

# Hardy inequalities, Rellich inequalities and local Dirichlet forms

Derek W. Robinson<sup>†</sup>

January 2017

**Abstract** First the Hardy and Rellich inequalities are defined for the submarkovian operator associated with a local Dirichlet form. Secondly, two general conditions are derived which are sufficient to deduce the Rellich inequality from the Hardy inequality. In addition the Rellich constant is calculated from the Hardy constant. Thirdly, we establish that the criteria for the Rellich inequality are verified for a large class of weighted second-order operators on a domain  $\Omega \subseteq \mathbf{R}^d$ . The weighting near the boundary  $\partial\Omega$  can be different from the weighting at infinity. Finally these results are applied to weighted second-order operators on  $\mathbf{R}^d \setminus \{0\}$  and to a general class of operators of Grushin type.

AMS Subject Classification: 31C25, 47D07, 39B62.

<sup>†</sup> Mathematical Sciences Institute (CMA)  
Australian National University  
Canberra, ACT 0200  
Australia  
derek.robinson@anu.edu.au

# 1 Introduction

Our intent is twofold. First we analyze Hardy and Rellich inequalities in the general framework of local Dirichlet forms. Secondly we apply the analysis to a large class of divergence-form elliptic operators on domains of  $\mathbf{R}^d$ . Our principal results give verifiable criteria that allow the deduction of the Rellich inequality from the Hardy inequality.

There is an enormous literature concerning variants of the Hardy inequality and their applications but somewhat less for the Rellich inequality. We refer to the book by Balinsky, Evans and Lewis [BEL15] and the thesis by Ward [War14] for background information and an indication of the relevant literature. In order to explain our results we first establish some notation and recall some basic elements of the theory of Dirichlet forms. We mostly adopt the definitions and terminology of Bouleau and Hirsch [BH91] (see also [FOT94]). Subsequently we turn to the examination of elliptic operators on domains of Euclidean space.

Let  $X$  denote a locally compact  $\sigma$ -compact metric space and  $\mu$  a positive Radon measure with  $\text{supp } \mu = X$ . The corresponding real  $L_p$ -spaces are denoted by  $L_p(X)$ . Let  $\mathcal{E}$  denote a Dirichlet form with domain  $D(\mathcal{E})$  on  $L_2(X)$  and  $H$  the self-adjoint submarkovian operator canonically associated with  $\mathcal{E}$ . Set  $B(\mathcal{E}) = D(\mathcal{E}) \cap L_\infty(X)$ . Then  $B(\mathcal{E})$  is an algebra and a core of  $D(\mathcal{E})$ . Further let  $B_{\text{loc}}(\mathcal{E})$  denote the corresponding space of functions which are locally in  $B(\mathcal{E})$ , i.e. the space of  $\mu$ -measurable functions  $\psi$  such that for every compact subset  $K$  of  $X$  there is a  $\hat{\psi} \in B(\mathcal{E})$  with  $\psi|_K = \hat{\psi}|_K$ . Next let  $B_c(\mathcal{E})$  denote the subalgebra of  $B(\mathcal{E})$  formed by the functions with compact support and set  $C_c(\mathcal{E}) = B_c(\mathcal{E}) \cap C(X)$ . We assume that  $C_c(\mathcal{E})$  is dense in  $C_0(X)$ , the space of continuous functions over  $X$  which vanish at infinity, with respect to the supremum norm and that it is also dense in  $B_c(\mathcal{E})$  with respect to the  $D(\mathcal{E})$ -graph norm, i.e. the norm  $\varphi \in D(\mathcal{E}) \mapsto \|\varphi\|_{D(\mathcal{E})} = (\mathcal{E}(\varphi) + \|\varphi\|_2^2)^{1/2}$ . In addition we assume that  $\mathcal{E}$  is local in the sense of Bouleau and Hirsch [BH91], Section I.1.5. In particular  $\mathcal{E}$  is local if  $\mathcal{E}(\varphi, \psi) = 0$  for all  $\varphi, \psi \in D(\mathcal{E})$  for which there is an  $a \in \mathbf{R}$  such that  $(\varphi + a\mathbb{1})\psi = 0$ . (A slightly more specific property is introduced in [FOT94], Section 1.1 and is referred to as strong locality.) Finally, for each positive  $\xi \in B(\mathcal{E})$  we define the truncated form  $\mathcal{E}_\xi$  by  $D(\mathcal{E}_\xi) = B(\mathcal{E})$  and

$$\mathcal{E}_\xi(\varphi) = \mathcal{E}(\varphi, \xi\varphi) - 2^{-1}\mathcal{E}(\xi, \varphi^2) \quad (1)$$

for all  $\varphi \in B(\mathcal{E})$ . The truncated forms satisfy the Markovian properties characteristic of Dirichlet forms but are not necessarily closed. Moreover,  $\xi \mapsto \mathcal{E}_\xi(\varphi)$  is a positive linear functional for each  $\varphi \in B(\mathcal{E})$  and  $\mathcal{E}_\xi(\varphi) \leq \|\xi\|_\infty \mathcal{E}(\varphi)$  for all  $\varphi \in B(\mathcal{E})$  (see [BH91], Proposition 4.1.1 for these properties). Consequently, for each  $\varphi \in B(\mathcal{E})$ , the function  $\xi \mapsto \mathcal{E}_\xi(\varphi)$  extends by continuity to  $C_0(X)$ . Then there is a positive Radon measure  $\mu_\varphi$ , the energy measure, such that  $\mu_\varphi(\xi) = \mathcal{E}_\xi(\varphi)$  for all  $\xi \in C_0(X)$ . Note that if  $\xi \in B(\mathcal{E})$  has compact support one can also define  $\mathcal{E}_\xi$  on  $B_{\text{loc}}(\mathcal{E}) \cap L_\infty(X)$  by setting  $\mathcal{E}_\xi(\varphi) = \mathcal{E}_\xi(\hat{\varphi})$  where  $\hat{\varphi} \in B(\mathcal{E})$  is such that  $\hat{\varphi}|_{\text{supp } \xi} = \varphi|_{\text{supp } \xi}$ . The definition is consistent by locality.

Next let  $\eta \in B_{\text{loc}}(\mathcal{E})$ . Then the Dirichlet form  $\mathcal{E}$  is defined to satisfy the  $\eta$ -Hardy inequality if  $\eta D(\mathcal{E}) \subseteq L_2(X)$  and

$$\mathcal{E}(\varphi) \geq (\eta\varphi, \eta\varphi) \quad (2)$$

for all  $\varphi \in D(\mathcal{E})$ . Since this condition is invariant under the map  $\eta \mapsto |\eta|$  one may always assume that  $\eta$  is positive. Similarly,  $\mathcal{E}$  is defined to satisfy the  $\eta$ -Rellich inequality if

$\eta^2 D(H) \subseteq L_2(X)$  and there is a  $\sigma > 0$  such that

$$(H\varphi, H\varphi) \geq \sigma (\eta^2 \varphi, \eta^2 \varphi) \quad (3)$$

for all  $\varphi \in D(H)$ .

Our main result, which is proved in Section 2, establishes conditions which ensure that the  $\eta$ -Rellich inequality follows from the  $\eta$ -Hardy inequality.

**Theorem 1.1** *Assume  $\mathcal{E}$  is a local Dirichlet form and  $\eta \in B_{\text{loc}}(\mathcal{E})$  is positive. Further assume*

- I.  $\mathcal{E}$  satisfies the  $\eta$ -Hardy inequality (2),
- II. there is a  $\gamma \in \langle 0, 1 \rangle$  such that  $\mathcal{E}_{\varphi^2}(\eta) \leq \gamma (\eta^2 \varphi, \eta^2 \varphi)$  for all  $\varphi \in B_c(\mathcal{E})$ ,
- III. there exists a net  $\{\rho_\alpha\}$  with  $\rho_\alpha \in B_c(\mathcal{E})$  such that  $0 \leq \rho_\alpha \leq 1$ ,

$$\lim_{\alpha} (\varphi, \rho_\alpha \varphi) = (\varphi, \varphi) \quad \text{and} \quad \lim_{\alpha} \mathcal{E}_{\varphi^2}(\rho_\alpha) = 0 \quad (4)$$

for all  $\varphi \in B(\mathcal{E})$ .

It then follows that  $\mathcal{E}$  satisfies the  $\eta$ -Rellich inequality (3) with  $\sigma = (1 - \gamma)^2$ .

The Hardy inequality (2) is the quadratic form expression of the ordering  $H \geq \eta^2$  of the self-adjoint operators  $H$  and  $\eta^2$  where the latter is interpreted as a multiplication operator. Similarly, the Rellich inequality (3) corresponds to the order relation  $H^2 \geq \sigma \eta^4$ . Note that the order relation is, however, not generally respected by taking squares unless the operators commute. This is the role played by Condition II; it imposes restrictions on the commutativity of  $H$  and  $\eta$ . The condition can be rephrased in terms of the energy measures  $\mu_\varphi$  and it is most transparent if these measures are absolutely continuous with respect to  $\mu$ . The corresponding Radon–Nikodym derivatives  $\Gamma: \varphi \in B(\mathcal{E}) \mapsto \Gamma(\varphi) \in L_1(\Omega; \mu)$ , which are usually referred to as the *carré du champ*, define a positive quadratic form whose basic properties are developed in [BH91] Section I.1.4. Then one has

$$\mathcal{E}_{\varphi^2}(\eta) = \int_X d\mu \Gamma(\eta) \varphi^2 = (\varphi, \Gamma(\eta) \varphi)$$

for all  $\varphi \in B_c(\mathcal{E})$ . Thus Condition II is equivalent to the bounds

$$0 \leq \Gamma(\eta) \leq \gamma \eta^4.$$

But in applications to elliptic operators  $\Gamma(\eta)$  is a measure of non-commutation. For example, the *carré du champ* corresponding to the Laplacian  $\Delta$  on  $\mathbf{R}^d$  is given by  $\Gamma(\eta) = |\nabla \eta|^2$  and formally  $|\nabla \eta|^2 = -2^{-1} [[\Delta, \eta], \eta]$ . Thus in this case Condition II leads to the bounds

$$0 \leq -2^{-1} [[\Delta, \eta], \eta] \leq \gamma \eta^4$$

on the double commutator of  $\Delta$  and  $\eta$ . This restriction on the commutation is the essential content of Condition II of the theorem. Double commutator estimates of this type occur in many disparate areas of mathematical physics and analysis, e.g. in quantum field theory, [GJ72] [GJ81] Section 19.4, [RS75] Section X.5, operator theory, [Far75] Section II.12, [DS83], elliptic equations [Agm82], elliptic regularity [FP83] [ER09] [RS11], etc.

Condition III of the theorem, which is independent of the Hardy–Rellich function  $\eta$ , corresponds to the existence of a special form of approximate identity  $\{\rho_\alpha\}$  on  $L_2(X)$ . Although it is not evident that an approximation of this type exists in general we do establish that it exists for a large class of divergence-form elliptic operators on a general domain of  $\mathbf{R}^d$  if the operators satisfy an appropriate Hardy inequality. To describe our results in the latter context we need some additional terminology.

Let  $\Omega$  be a domain in  $\mathbf{R}^d$ , i.e. a connected open subset, with boundary  $\partial\Omega$  and equipped with the Euclidean metric  $d(\cdot; \cdot)$ . Further let  $x \in \Omega \mapsto d_\Omega(x) \in \langle 0, \infty \rangle$  denote the Euclidean distance to the boundary, i.e.  $d_\Omega(x) = \inf_{y \in \Omega^c} d(x; y)$ . If  $c$  is a strictly positive function on the half line  $\langle 0, \infty \rangle$  we define  $c_\Omega$  by  $c_\Omega = c \circ d_\Omega$ . Then we define a Dirichlet form  $h$  on  $L_2(\Omega)$  as the closure of the form

$$\varphi \in C_c^\infty(\Omega) \mapsto h(\varphi) = \sum_{k=1}^d (\partial_k \varphi, c_\Omega \partial_k \varphi) . \quad (5)$$

The form is closable, because  $c$  is strictly positive, and the closed form is automatically a local Dirichlet form (see, for example, [MR92] Section II.2.b). Moreover, the form has a *carré du champ*  $\Gamma$  given by  $\Gamma(\varphi) = c_\Omega |\nabla \varphi|^2$ . The submarkovian operator  $H$  corresponding to the form  $h$  can be interpreted as the elliptic operator  $-\sum_{k=1}^d \partial_k c_\Omega \partial_k$  with Dirichlet boundary conditions. In the context of the Hardy–Rellich inequality the choice  $c(s) = s^\delta$  is conventional but we will consider a broader class of coefficients and operators.

Our second principal result is essentially a corollary of Theorem 1.1.

**Theorem 1.2** *Let  $c(s) = s^\delta (1+s)^{\delta'-\delta}$  with  $\delta, \delta' \geq 0$  and set*

$$\nu = \sup\{|1 - t/2|^2 : \delta \wedge \delta' \leq t \leq \delta \vee \delta'\} .$$

*Assume that the Dirichlet form (5) on  $L_2(\Omega)$  corresponding to  $c$  satisfies the Hardy inequality*

$$h(\varphi) \geq a_1 (c_\Omega^{1/2} d_\Omega^{-1} \varphi, c_\Omega^{1/2} d_\Omega^{-1} \varphi) \quad (6)$$

*for all  $\varphi \in D(h)$  with  $a_1 > 0$ .*

*If  $\nu < a_1$  then  $H$  satisfies the Rellich inequality*

$$(H\varphi, H\varphi) \geq a_2 (c_\Omega d_\Omega^{-2} \varphi, c_\Omega d_\Omega^{-2} \varphi) \quad (7)$$

*for all  $\varphi \in D(H)$  with  $a_2 = (a_1 - \nu)^2$ .*

Our choice of the weight function  $c$  in Theorem 1.2 is dictated by its asymptotic behaviour. The parameters  $\delta$  and  $\delta'$  govern the growth properties of  $c_\Omega$  near the boundary and at infinity, respectively. In particular one has  $\lim_{s \rightarrow 0} c(s) s^{-\delta} = 1$  and  $\lim_{s \rightarrow \infty} c(s) s^{-\delta'} = 1$ .

Note that Condition I of Theorem 1.1 is satisfied with  $\eta^2 = a_1 c_\Omega d_\Omega^{-2}$  by the assumption that  $h$  satisfies the Hardy inequality (6). Moreover Condition II of the earlier theorem is not difficult to verify by direct calculation using the properties of  $c$ . But the verification of the third condition of Theorem 1.1 is more difficult. We achieve this by adaptation of an argument introduced by Agmon [Agm82] in his analysis of the exponential decay of solutions of second-order elliptic equations. Agmon's arguments have earlier been used by Grillo [Gri03] to discuss Hardy and Rellich inequalities for operators constructed as sums of squares of vector fields.

Finally we note that the conclusions of the theorems are established for all functions in the domain  $D(H)$  of the submarkovian operator  $H$ . Many derivations of the Rellich inequality are only valid on a core  $D$  of the corresponding form  $\mathcal{E}$  but not for a core of  $H$ . In particular if  $H_0$  is a symmetric elliptic operator on a domain  $\Omega \subseteq \mathbf{R}^d$  it is commonplace to use  $D = C_c^\infty(\Omega)$  (see, for example, [BEL15], Chapter 6). Then the Dirichlet form  $\mathcal{E}$  obtained by closure of  $\varphi \in C_c^\infty(\Omega) \mapsto (\varphi, H_0\varphi)$  determines the self-adjoint Friedrichs extension  $H_F$  of  $H_0$  but the closure of the form  $\varphi \in C_c^\infty(\Omega) \mapsto (H_0\varphi, H_0\varphi)$  determines the Friedrichs extension  $(H_0^2)_F$  of  $H_0^2$  which usually differs from  $H_F^2$ . In general one has  $(H_0^2)_F \geq H_F^2$  with equality if and only if  $H_0$  is essentially self-adjoint.

## 2 Locality estimates

In this section we give the proof of Theorem 1.1. It is based on several identities and estimates which are a direct result of the structure of the Dirichlet form  $\mathcal{E}$  and the locality condition. We begin by collecting some specific results of relevance to the proof.

The locality property can be exploited by use of Anderssen's representation theorem [And75] (see also [Rot76]) which is reformulated in [BH91], Theorem I.5.2.1. We will use the formulation given in [ERSZ06] (see also [AH05]).

**Proposition 2.1** *Let  $\mathcal{E}$  be a local Dirichlet form on  $L_2(X)$  and  $\varphi_1, \dots, \varphi_n \in B(\mathcal{E})$ . Then there exists a unique real Radon measure  $\sigma_{ij}^{(\varphi_1, \dots, \varphi_n)}$  on  $\mathbf{R}^n$  such that  $\sigma_{ij}^{(\varphi_1, \dots, \varphi_n)} = \sigma_{ji}^{(\varphi_1, \dots, \varphi_n)}$  for all  $i, j \in \{1, \dots, n\}$ ,  $\sum_{i,j=1}^n \xi_i \xi_j \sigma_{ij} \geq 0$  for all  $\xi_i, \xi_j \in \mathbf{R}$  and*

$$\mathcal{E}(F_0(\varphi_1, \dots, \varphi_n), G_0(\varphi_1, \dots, \varphi_n)) = \sum_{i,j=1}^n \int_{\mathbf{R}^n} d\sigma_{ij}^{(\varphi_1, \dots, \varphi_n)} \frac{\partial F}{\partial x_i} \frac{\partial G}{\partial x_j} \quad (8)$$

for all  $F, G \in C^1(\mathbf{R}^n)$  where  $F_0 = F - F(0)$  and  $G_0 = G - G(0)$ . Let  $K$  be a compact subset of  $\mathbf{R}^n$  such that  $(\varphi_1(x), \dots, \varphi_n(x)) \in K$  for almost every  $x \in X$ . Then  $\text{supp } \sigma_{ij}^{(\varphi_1, \dots, \varphi_n)} \subseteq K$  for all  $i, j \in \{1, \dots, n\}$ . In particular, if  $i \in \{1, \dots, n\}$  then  $\sigma_{ii}^{(\varphi_1, \dots, \varphi_n)}$  is a finite (positive) measure.

One immediate implication of Proposition 2.1 is the Leibniz rule, or derivation property,

$$\mathcal{E}_\chi(\varphi_1\varphi_2, \psi_1\psi_2) = \mathcal{E}_{\varphi_1\psi_1\chi}(\varphi_2, \psi_2) + \mathcal{E}_{\varphi_1\psi_2\chi}(\varphi_2, \psi_1) + \mathcal{E}_{\varphi_2\psi_1\chi}(\varphi_1, \psi_2) + \mathcal{E}_{\varphi_2\psi_2\chi}(\varphi_1, \psi_1)$$

for the bilinear form  $\mathcal{E}_\chi(\varphi, \psi)$  related to the truncated form  $\mathcal{E}_\chi$  by polarization and a similar identity for the bilinear form  $\mathcal{E}(\varphi, \psi)$  associated with  $\mathcal{E}$ . (The latter identity is formally obtained from the former by setting  $\chi = \mathbf{1}_X$  and  $\mathcal{E}_{\mathbf{1}_X} = \mathcal{E}$ .)

Our next application of the proposition is a key identity related to the estimate given by Condition II of Theorem 1.1.

**Lemma 2.2** *If  $\varphi, \chi \in B(\mathcal{E})$  then*

$$\mathcal{E}_{\varphi^2}(\chi) = \mathcal{E}(\chi\varphi) - \mathcal{E}(\varphi, \chi^2\varphi). \quad (9)$$

**Proof** Let  $\sigma_{ij}$  denote the representing measure of Proposition 2.1 corresponding to the pair  $\varphi, \chi$ . Then

$$\mathcal{E}(\chi \varphi) = \int d\sigma_{11}(x_1, x_2) x_2^2 + 2 \int d\sigma_{12}(x_1, x_2) x_1 x_2 + \int d\sigma_{22}(x_1, x_2) x_1^2 \quad (10)$$

and

$$\mathcal{E}(\varphi, \chi^2 \varphi) = \int d\sigma_{11}(x_1, x_2) x_2^2 + 2 \int d\sigma_{12}(x_1, x_2) x_1 x_2 .$$

Therefore

$$\mathcal{E}(\chi \varphi) - \mathcal{E}(\varphi, \chi^2 \varphi) = \int d\sigma_{22}(x_1, x_2) x_1^2 . \quad (11)$$

Similarly, one calculates that

$$\mathcal{E}_{\varphi^2}(\chi) = \mathcal{E}(\chi, \varphi^2 \chi) - 2^{-1} \mathcal{E}(\varphi^2, \chi^2) = \int d\sigma_{22}(x_1, x_2) x_1^2 \quad (12)$$

Then (9) follows directly from (11) and (12).  $\square$

The relevance of the identity (9) is that it formally identifies the energy measure corresponding to  $\mathcal{E}$  with a double commutator. To illustrate this assume  $\mathcal{E}$  has a *carré du champ*  $\Gamma$ . Then it follows that

$$\mathcal{E}_{\varphi^2}(\psi) = \int_X d\mu \varphi^2 \Gamma(\psi) = (\varphi, \Gamma(\psi) \varphi)$$

for  $\varphi, \psi \in B(\mathcal{E})$ . Therefore (9) gives the identification

$$-2(\varphi, \Gamma(\psi) \varphi) = \mathcal{E}(\varphi, \psi^2 \varphi) - 2\mathcal{E}(\psi \varphi, \psi \varphi) + \mathcal{E}(\psi^2 \varphi, \varphi)$$

for all  $\varphi, \psi \in B(\mathcal{E})$ . But if  $\psi \varphi \in D(H)$  for each  $\varphi \in D(H)$  then

$$-2(\varphi, \Gamma(\psi) \varphi) = (H\varphi, \psi^2 \varphi) - 2(\psi \varphi, H\psi \varphi) + (\psi^2 \varphi, H\varphi)$$

which is equivalent to the identification

$$\Gamma(\psi) = -2^{-1}[[H, \psi], \psi]$$

analogous to the situation for the Laplacian discussed in the introduction.

Next we need the following estimate.

**Lemma 2.3** *If  $\varphi, \chi \in B(\mathcal{E})$  with  $\chi \geq 0$  then*

$$\mathcal{E}_{\varphi^2}(\chi(1 + \beta\chi)^{-1}) \leq \mathcal{E}_{(1+\beta\chi)^{-2}\varphi^2}(\chi)$$

for all  $\beta \geq 0$ .

**Proof** Again let  $\sigma_{ij}$  denote the representing measure corresponding to the pair  $\varphi, \chi$ . Then one calculates that

$$\begin{aligned} \mathcal{E}_{\varphi^2}(\chi(1 + \beta\chi)^{-1}) &= \mathcal{E}(\chi(1 + \beta\chi)^{-1}, \varphi^2 \chi(1 + \beta\chi)^{-1}) - 2^{-1} \mathcal{E}(\chi^2(1 + \beta\chi)^{-2}, \varphi^2) \\ &= \int d\sigma_{22}(x_1, x_2) x_1^2 (1 + \beta x_2)^{-4} \end{aligned}$$

since the terms corresponding to  $\sigma_{12}$  cancel. Similarly

$$\begin{aligned}\mathcal{E}_{(1+\beta\chi)^{-2}\varphi^2}(\chi) &= \mathcal{E}(\chi, (1+\beta\chi)^{-2}\varphi^2\chi) - 2^{-1}\mathcal{E}(\chi^2, (1+\beta\chi)^{-2}\varphi^2) \\ &= \int d\sigma_{22}(x_1, x_2) x_1^2 (1+\beta x_2)^{-2}\end{aligned}$$

because the cross terms again cancel. (Since  $\chi \geq 0$  the  $x_2$ -integration is over the positive half axis.) Therefore the statement of the lemma follows immediately.  $\square$

A locality estimate also gives the following bounds.

**Lemma 2.4** *If  $\varphi, \chi \in B(\mathcal{E})$  then*

$$\mathcal{E}(\chi\varphi) \leq (1+\delta)\mathcal{E}_{\chi^2}(\varphi) + (1+\delta^{-1})\mathcal{E}_{\varphi^2}(\chi)$$

*for all  $\delta > 0$ . If, in addition,  $\psi \in B(\mathcal{E})$  then*

$$\mathcal{E}_{\psi^2}(\chi\varphi) \leq (1+\delta)\mathcal{E}_{\psi^2\chi^2}(\varphi) + (1+\delta^{-1})\mathcal{E}_{\psi^2\varphi^2}(\chi)$$

*for all  $\delta > 0$*

**Proof** It follows from (10) and the Cauchy–Schwarz inequality for the measure  $\sigma_{ij}$  that

$$\mathcal{E}(\chi\varphi) \leq (1+\delta^{-1}) \int d\sigma_{11}(x_1, x_2) x_2^2 + (1+\delta) \int d\sigma_{22}(x_1, x_2) x_1^2.$$

But the second integral on the right is equal to  $\mathcal{E}_{\chi^2}(\varphi)$  by (12). Moreover, by interchanging  $\chi$  and  $\varphi$  one can identify the first integral with  $\mathcal{E}_{\varphi^2}(\chi)$ . The first statement of the lemma follows by combination of these observations. The second statement follows by a similar calculation.  $\square$

Next one has the following identity.

**Lemma 2.5** *If  $\varphi, \chi \in B(\mathcal{E})$  then*

$$\mathcal{E}_{\varphi^2}(\chi^2) = 4\mathcal{E}_{\chi^2\varphi^2}(\chi).$$

The proof follows by a similar calculation to the derivation of (12). Alternatively it follows directly from the Leibniz rule.

Finally we consider approximation of functions in the domain  $D(H)$  of the submarkovian operator  $H$ . One has  $D(H) \subseteq D(\mathcal{E})$  but it is convenient to establish an explicit approximation, in the  $D(\mathcal{E})$ -graph norm, of functions in  $D(H)$  by functions in  $B(\mathcal{E})$ .

**Lemma 2.6** *If  $\varphi \in D(H)$  and  $\varepsilon > 0$  set  $\varphi_\varepsilon = \varphi(1+\varepsilon\varphi^2)^{-1/2}$ . Then  $\varphi_\varepsilon \in B(\mathcal{E})$ ,  $\|\varphi_\varepsilon\|_2 \leq \|\varphi\|_2$ ,  $\mathcal{E}(\varphi_\varepsilon) \leq \mathcal{E}(\varphi)$  and  $\|\varphi - \varphi_\varepsilon\|_{D(\mathcal{E})} \rightarrow 0$  as  $\varepsilon \rightarrow 0$ . Moreover,  $\mathcal{E}_\xi(\varphi_\varepsilon) \leq \mathcal{E}_\xi(\varphi)$  for all  $\xi \in B(\mathcal{E})_+$  and  $\mathcal{E}_\xi(\varphi_\varepsilon) \rightarrow \mathcal{E}_\xi(\varphi)$  as  $\varepsilon \rightarrow 0$ .*

**Proof** First the boundedness property  $\|\varphi_\varepsilon\|_2 \leq \|\varphi\|_2$  is evident. Secondly

$$\|\varphi - \varphi_\varepsilon\|_2^2 = \|\varphi(1 - (1+\varepsilon\varphi^2)^{-1/2})\|_2^2 = \int_\Omega d\mu \varphi^2 (1 - (1+\varepsilon\varphi^2)^{-1/2})^2 \leq \|\varphi\|_2^2.$$

Therefore  $\|\varphi - \varphi_\varepsilon\|_2 \rightarrow 0$  as  $\varepsilon \rightarrow 0$  by the Lebesgue dominated convergence theorem.

Thirdly, the map  $\varphi \mapsto \varphi_\varepsilon$  is a normal contraction so  $\mathcal{E}(\varphi_\varepsilon) \leq \mathcal{E}(\varphi)$  and  $\mathcal{E}_\xi(\varphi_\varepsilon) \leq \mathcal{E}_\xi(\varphi)$  by the Markovian property of the form  $\mathcal{E}$  and the associated truncated functions. The remaining convergence statements follow from the Anderssen representation, e.g. if  $\sigma^\varphi$  denotes the positive measure corresponding to  $\varphi \in D(\mathcal{E})$  then

$$\mathcal{E}(\varphi - \varphi_\varepsilon) = \int d\sigma^\varphi(x) (1 - (1 + \varepsilon x^2)^{-3/2})^2 \leq \mathcal{E}(\varphi) .$$

Therefore  $\mathcal{E}(\varphi - \varphi_\varepsilon) \rightarrow 0$  as  $\varepsilon \rightarrow 0$  by another application of dominated convergence. Consequently,  $\varphi_\varepsilon \rightarrow \varphi$  as  $\varepsilon \rightarrow 0$  with respect to the  $D(\mathcal{E})$ -graph norm.  $\square$

At this point we are prepared to prove Theorem 1.1. The proof of is in two steps. First we establish the conclusion for functions  $\varphi \in D(H) \cap B_c(\mathcal{E})$ . Secondly, we extend the result to all  $\varphi \in D(H)$  by approximation. The first step is straightforward but the second step is more complicated. It involves simultaneous approximation of the Hardy function  $\eta$  by bounded functions of compact support and  $\varphi$  by bounded functions in the form domain.

**Proof of Theorem 1.1** Let  $\eta \in B_{\text{loc}}(\mathcal{E})$  be such that the  $\eta$ -Hardy inequality  $\mathcal{E}(\varphi) \geq (\eta\varphi, \eta\varphi)$  is satisfied for all  $\varphi \in B_c(\mathcal{E})$ . Since  $B(\mathcal{E})$  is an algebra it follows that  $B_{\text{loc}}(\mathcal{E})B_c(\mathcal{E}) \subseteq B_c(\mathcal{E})$ . Therefore  $\mathcal{E}(\eta\varphi) \geq (\eta^2\varphi, \eta^2\varphi)$  for all  $\varphi \in B_c(\mathcal{E})$ . Then it follows from Lemma 2.2 that

$$\mathcal{E}(\varphi, \eta^2\varphi) \geq (\eta^2\varphi, \eta^2\varphi) - \mathcal{E}_{\varphi^2}(\eta)$$

for all  $\varphi \in B_c(\mathcal{E})$ . But  $\mathcal{E}_{\varphi^2}(\eta) \leq \gamma (\eta^2\varphi, \eta^2\varphi)$  with  $\gamma \in \langle 0, 1 \rangle$  by Condition II of the theorem. Therefore

$$\mathcal{E}(\varphi, \eta^2\varphi) \geq (1 - \gamma)(\eta^2\varphi, \eta^2\varphi)$$

for all  $\varphi \in B_c(\mathcal{E})$ . If, however,  $\varphi \in D(H) \cap B_c(\mathcal{E})$  then  $|\mathcal{E}(\varphi, \eta^2\varphi)| \leq \|H\varphi\|_2 \|\eta^2\varphi\|_2$  and one deduces that

$$\|H\varphi\|_2 \|\eta^2\varphi\|_2 \geq (1 - \gamma)(\eta^2\varphi, \eta^2\varphi) = (1 - \gamma)\|\eta^2\varphi\|_2^2 .$$

Since  $\gamma \in \langle 0, 1 \rangle$  one can divide by  $\|\eta^2\varphi\|_2$  and square to obtain the Rellich inequality

$$(H\varphi, H\varphi) = \|H\varphi\|_2^2 \geq (1 - \gamma)^2(\eta^2\varphi, \eta^2\varphi)$$

for all  $\varphi \in D(H) \cap B_c(\mathcal{E})$ .

The derivation of the inequality for general  $\varphi \in D(H)$  follows a similar line of reasoning but is essentially more complicated. As a preliminary note that the definition of the  $\eta$ -Hardy inequality includes the condition  $\eta D(\mathcal{E}) \subseteq L_2(X)$  but the Rellich inequality requires that  $\eta^2 D(H) \subseteq L_2(X)$ . Since  $\eta \in B_{\text{loc}}(\mathcal{E})$  and  $B(\mathcal{E})$  is an algebra one always has the weaker properties  $\eta B_c(\mathcal{E}) \subseteq L_2(X)$  and  $\eta^2 B_c(\mathcal{E}) \subseteq L_2(X)$ . These relations do not depend on the validity of the Hardy or Rellich inequalities. But the conditions  $\eta D(\mathcal{E}) \subseteq L_2(X)$  and  $\eta^2 D(H) \subseteq L_2(X)$  are much more stringent. This is the reason that the proof for general  $\varphi \in D(H)$  is complicated by additional approximation arguments.

Fix  $\varphi \in D(H)$  and set  $\varphi_\varepsilon = \varphi(1 + \varepsilon \varphi^2)^{-1/2}$  with  $\varepsilon > 0$ . Then  $\varphi_\varepsilon \in B(\mathcal{E})$ , by Lemma 2.6. Further set  $\eta_{\alpha,\beta} = \rho_\alpha \eta_\beta$  where the  $\rho_\alpha \in B_c(\mathcal{E})$  satisfy Condition III of the theorem and



$\eta_\beta = \eta(1 + \beta\eta)^{-1}$  with  $\beta > 0$ . Then  $\eta_\beta \in B(\mathcal{E})$ ,  $\eta_{\alpha,\beta} \in B_c(\mathcal{E})$  and  $0 \leq \eta_{\alpha,\beta} \leq \eta_\beta \leq \eta$ . Moreover,

$$\begin{aligned} (H\varphi, \eta_{\alpha,\beta}^2 \varphi_\varepsilon) &= \mathcal{E}(\varphi, \eta_{\alpha,\beta}^2 \varphi_\varepsilon) \\ &= \mathcal{E}(\varphi_\varepsilon, \eta_{\alpha,\beta}^2 \varphi_\varepsilon) + \mathcal{E}(\varphi - \varphi_\varepsilon, \eta_{\alpha,\beta}^2 \varphi_\varepsilon) \\ &= \mathcal{E}(\eta_{\alpha,\beta} \varphi_\varepsilon) - \mathcal{E}_{\varphi_\varepsilon^2}(\eta_{\alpha,\beta}) + \mathcal{E}(\varphi - \varphi_\varepsilon, \eta_{\alpha,\beta}^2 \varphi_\varepsilon) \end{aligned} \quad (13)$$

where the third step follows from Lemma 2.2. In addition  $\mathcal{E}(\eta_{\alpha,\beta} \varphi_\varepsilon) \geq (\eta \eta_{\alpha,\beta} \varphi_\varepsilon, \eta \eta_{\alpha,\beta} \varphi_\varepsilon)$  by the  $\eta$ -Hardy inequality, Condition I of the theorem, applied to  $\eta_{\alpha,\beta} \varphi_\varepsilon$ . Therefore

$$\begin{aligned} (H\varphi, \eta_{\alpha,\beta}^2 \varphi_\varepsilon) &\geq (\eta \eta_{\alpha,\beta} \varphi_\varepsilon, \eta \eta_{\alpha,\beta} \varphi_\varepsilon) - \mathcal{E}_{\varphi_\varepsilon^2}(\eta_{\alpha,\beta}) + \mathcal{E}(\varphi - \varphi_\varepsilon, \eta_{\alpha,\beta}^2 \varphi_\varepsilon) \\ &\geq (\eta_{\alpha,\beta} \eta \varphi_\varepsilon, \eta_{\alpha,\beta} \eta \varphi_\varepsilon) - \mathcal{E}_{\varphi_\varepsilon^2}(\eta_{\alpha,\beta}) - \mathcal{E}(\varphi - \varphi_\varepsilon)^{1/2} \mathcal{E}(\eta_{\alpha,\beta}^2 \varphi_\varepsilon)^{1/2} \end{aligned} \quad (14)$$

for all  $\varphi \in D(H)$ , all  $\beta, \varepsilon > 0$  and all  $\alpha$ . But

$$|(H\varphi, \eta_{\alpha,\beta}^2 \varphi_\varepsilon)| \leq \|H\varphi\|_2 \|\eta_\beta^2 \varphi_\varepsilon\|_2 \leq \|H\varphi\|_2 \|\eta_\beta \eta \varphi_\varepsilon\|_2$$

since  $\eta \varphi_\varepsilon \in L_2(X)$  by the  $\eta$ -Hardy inequality. Combining this estimate with (14) and taking a limit over  $\alpha$  then gives

$$\begin{aligned} \|H\varphi\|_2 \|\eta_\beta \eta \varphi_\varepsilon\|_2 &\geq \|\eta_\beta \eta \varphi_\varepsilon\|_2^2 - \limsup_{\alpha} \mathcal{E}_{\varphi_\varepsilon^2}(\eta_{\alpha,\beta}) \\ &\quad - \mathcal{E}(\varphi - \varphi_\varepsilon)^{1/2} \limsup_{\alpha} \mathcal{E}(\eta_{\alpha,\beta}^2 \varphi_\varepsilon)^{1/2} \end{aligned} \quad (15)$$

for all  $\beta, \varepsilon > 0$ . Here we have used Condition III to deduce that  $\eta_{\alpha,\beta}$  converges on  $L_2(X)$  to  $\eta_\beta$ . It is important at this point that  $\eta \varphi_\varepsilon \in L_2(X)$  by the  $\eta$ -Hardy inequality and  $\eta_\beta \in L_\infty(X)$ . Therefore  $\eta_\beta \eta \varphi_\varepsilon \in L_2(X)$ .

Next consider the second term on the right hand side of (15). Since  $\eta_{\alpha,\beta} = \rho_\alpha \eta_\beta$  it follows from Lemma 2.4 that

$$\mathcal{E}_{\varphi_\varepsilon^2}(\eta_{\alpha,\beta}) \leq (1 + \delta) \mathcal{E}_{\rho_\alpha^2 \varphi_\varepsilon^2}(\eta_\beta) + (1 + \delta^{-1}) \mathcal{E}_{\eta_\beta^2 \varphi_\varepsilon^2}(\rho_\alpha)$$

for all  $\delta > 0$ . Now we apply Lemma 2.3, with  $\varphi$  replaced by  $\rho_\alpha \varphi_\varepsilon$  and  $\chi$  replaced by  $\eta$ , and Condition II of the theorem, with  $\varphi$  replaced by  $(1 + \beta\eta)^{-1} \rho_\alpha \varphi_\varepsilon$ , to the first term. One finds

$$\begin{aligned} \mathcal{E}_{\rho_\alpha^2 \varphi_\varepsilon^2}(\eta_\beta) &\leq \mathcal{E}_{(1+\beta\eta)^{-2} \rho_\alpha^2 \varphi_\varepsilon^2}(\eta) \\ &\leq \gamma (\eta^2 (1 + \beta\eta)^{-1} \rho_\alpha \varphi_\varepsilon, \eta^2 (1 + \beta\eta)^{-1} \rho_\alpha \varphi_\varepsilon) \leq \gamma \|\eta_\beta \eta \varphi_\varepsilon\|_2^2 \end{aligned}$$

where we again have  $\eta \varphi_\varepsilon \in L_2(X)$  by the  $\eta$ -Hardy inequality and  $\eta_\beta \in L_\infty(X)$ . Note that these steps are valid because  $\rho_\alpha$  has compact support. Therefore one now has

$$\mathcal{E}_{\varphi_\varepsilon^2}(\eta_{\alpha,\beta}) \leq (1 + \delta) \gamma \|\eta_\beta \eta \varphi_\varepsilon\|_2^2 + (1 + \delta^{-1}) \mathcal{E}_{\eta_\beta^2 \varphi_\varepsilon^2}(\rho_\alpha)$$

for all  $\delta > 0$ . Hence

$$\limsup_{\alpha} \mathcal{E}_{\varphi_\varepsilon^2}(\eta_{\alpha,\beta}) \leq (1 + \delta) \gamma \|\eta_\beta \eta \varphi_\varepsilon\|_2^2$$

for all  $\delta > 0$  since  $\limsup_{\alpha} \mathcal{E}_{\eta_{\beta}^2 \varphi_{\varepsilon}^2}(\rho_{\alpha}) = 0$  by Condition III of the theorem. Combining this estimate with (15) and taking the limit of  $\delta \rightarrow 0$  one then obtains the bounds

$$\|H\varphi\|_2 \|\eta_{\beta} \eta \varphi_{\varepsilon}\|_2 \geq (1 - \gamma) \|\eta_{\beta} \eta \varphi_{\varepsilon}\|_2^2 - \mathcal{E}(\varphi - \varphi_{\varepsilon})^{1/2} \limsup_{\alpha} \mathcal{E}(\eta_{\alpha,\beta}^2 \varphi_{\varepsilon})^{1/2} \quad (16)$$

for all  $\beta, \varepsilon > 0$ . Our next aim it to prove that  $\limsup_{\alpha} \mathcal{E}(\eta_{\alpha,\beta}^2 \varphi_{\varepsilon})$  is bounded uniformly in  $\varepsilon$ .

First consider  $\mathcal{E}(\eta_{\alpha,\beta} \varphi_{\varepsilon})$ . It follows that

$$\mathcal{E}(\eta_{\alpha,\beta} \varphi_{\varepsilon}) \leq 2 \mathcal{E}_{\eta_{\alpha,\beta}^2}(\varphi_{\varepsilon}) + 2 \mathcal{E}_{\varphi_{\varepsilon}^2}(\eta_{\alpha,\beta}) \leq 2 \|\eta_{\beta}\|_{\infty}^2 \mathcal{E}(\varphi_{\varepsilon}) + 2 \mathcal{E}_{\varphi_{\varepsilon}^2}(\eta_{\alpha,\beta})$$

where the first estimate uses Lemma 2.4 and the second uses  $\mathcal{E}_{\chi^2}(\psi) \leq \|\chi\|_{\infty}^2 \mathcal{E}(\psi)$  and  $\|\eta_{\alpha,\beta}\|_{\infty} \leq \|\eta_{\beta}\|_{\infty}$ . Then another application of Lemma 2.4 gives

$$\mathcal{E}_{\varphi_{\varepsilon}^2}(\eta_{\alpha,\beta}) \leq 2 \mathcal{E}_{\rho_{\alpha}^2 \varphi_{\varepsilon}^2}(\eta_{\beta}) + 2 \mathcal{E}_{\eta_{\beta}^2 \varphi_{\varepsilon}^2}(\rho_{\alpha}) .$$

Combining these bounds one obtains

$$\mathcal{E}(\eta_{\alpha,\beta} \varphi_{\varepsilon}) \leq 2 \|\eta_{\beta}\|_{\infty}^2 \mathcal{E}(\varphi_{\varepsilon}) + 4 \mathcal{E}_{\rho_{\alpha}^2 \varphi_{\varepsilon}^2}(\eta_{\beta}) + 4 \mathcal{E}_{\eta_{\beta}^2 \varphi_{\varepsilon}^2}(\rho_{\alpha}) .$$

But  $\mathcal{E}(\varphi_{\varepsilon}) \leq \mathcal{E}(\varphi)$  by Lemma 2.6. Moreover,

$$\mathcal{E}_{\rho_{\alpha}^2 \varphi_{\varepsilon}^2}(\eta_{\beta}) \leq \mathcal{E}_{(1+\beta\eta)^{-2} \rho_{\alpha}^2 \varphi_{\varepsilon}^2}(\eta) \leq \gamma (\eta_{\beta} \eta \rho_{\alpha} \varphi_{\varepsilon}, \eta_{\beta} \eta \rho_{\alpha} \varphi_{\varepsilon}) \leq \gamma \|\eta_{\beta} \eta \varphi\|_2^2 \quad (17)$$

by another application of Lemma 2.3 and Condition II of the theorem. Again this is valid since  $\rho_{\alpha}$  has compact support. Therefore combination of these last two estimates gives

$$\limsup_{\alpha} \mathcal{E}(\eta_{\alpha,\beta} \varphi_{\varepsilon}) \leq 2 \|\eta_{\beta}\|_{\infty}^2 \mathcal{E}(\varphi) + 4 \gamma \|\eta_{\beta} \eta \varphi\|_2^2$$

since  $\limsup_{\alpha} \mathcal{E}_{\eta_{\beta}^2 \varphi_{\varepsilon}^2}(\rho_{\alpha}) = 0$  by Condition III. Note that this last bound is uniform in  $\varepsilon$ .

Next by Lemma 2.4 one has

$$\mathcal{E}(\eta_{\alpha,\beta}^2 \varphi_{\varepsilon}) \leq 2 \mathcal{E}_{\eta_{\alpha,\beta}^2}(\eta_{\alpha,\beta} \varphi_{\varepsilon}) + 2 \mathcal{E}_{\eta_{\alpha,\beta}^2 \varphi_{\varepsilon}^2}(\eta_{\alpha,\beta}) \leq 2 \|\eta_{\beta}\|_{\infty}^2 \mathcal{E}(\eta_{\alpha,\beta} \varphi_{\varepsilon}) + 2 \mathcal{E}_{\eta_{\alpha,\beta}^2 \varphi_{\varepsilon}^2}(\eta_{\alpha,\beta})$$

and arguing as in the last paragraph

$$\mathcal{E}_{\eta_{\alpha,\beta}^2 \varphi_{\varepsilon}^2}(\eta_{\alpha,\beta}) \leq 2 \mathcal{E}_{\rho_{\alpha}^2 \eta_{\alpha,\beta}^2 \varphi_{\varepsilon}^2}(\eta_{\beta}) + 2 \mathcal{E}_{\eta_{\beta}^2 \eta_{\alpha,\beta}^2 \varphi_{\varepsilon}^2}(\rho_{\alpha}) \leq 2 \mathcal{E}_{\rho_{\alpha}^2 \eta_{\beta}^2 \varphi_{\varepsilon}^2}(\eta_{\beta}) + 2 \mathcal{E}_{\eta_{\beta}^4 \varphi_{\varepsilon}^2}(\rho_{\alpha}) .$$

Therefore combining these estimates one finds

$$\limsup_{\alpha} \mathcal{E}(\eta_{\alpha,\beta}^2 \varphi_{\varepsilon}) \leq 2 \|\eta_{\beta}\|_{\infty}^2 \limsup_{\alpha} \mathcal{E}(\eta_{\alpha,\beta} \varphi_{\varepsilon}) + 4 \limsup_{\alpha} \mathcal{E}_{\rho_{\alpha}^2 \eta_{\beta}^2 \varphi_{\varepsilon}^2}(\eta_{\beta}) .$$

The first term on the right is bounded uniformly in  $\varepsilon$  by the previous argument and

$$\mathcal{E}_{\rho_{\alpha}^2 \eta_{\beta}^2 \varphi_{\varepsilon}^2}(\eta_{\beta}) \leq \gamma (\eta_{\beta}^2 \eta \rho_{\alpha} \varphi_{\varepsilon}, \eta_{\beta}^2 \eta \rho_{\alpha} \varphi_{\varepsilon}) \leq \gamma \|\eta_{\beta}^2 \eta \varphi\|_2^2$$

by the estimate (17) with  $\varphi_{\varepsilon}$  replaced by  $\eta_{\beta} \varphi_{\varepsilon}$ . Therefore  $\limsup_{\alpha} \mathcal{E}(\eta_{\alpha,\beta}^2 \varphi_{\varepsilon})$  is bounded uniformly in  $\varepsilon$ .

Now one can take the limit  $\varepsilon \rightarrow 0$  in (16). Since  $\|\varphi - \varphi_\varepsilon\|_{D(\mathcal{E})} \rightarrow 0$  by Lemma 2.6, and  $\eta\varphi \in L_2(X)$  by the  $\eta$ -Hardy inequality, it follows that

$$\begin{aligned} \|H\varphi\|_2 \|\eta_\beta \eta \varphi\|_2 &= \lim_{\varepsilon \rightarrow 0} \|H\varphi\|_2 \|\eta_\beta \eta \varphi_\varepsilon\|_2 \\ &\geq (1 - \gamma) \limsup_{\varepsilon \rightarrow 0} \|\eta_\beta \eta \varphi_\varepsilon\|_2^2 = (1 - \gamma) \|\eta_\beta \eta \varphi\|_2^2. \end{aligned}$$

Therefore

$$\|H\varphi\|_2 \|\eta_\beta \eta \varphi\|_2 \geq (1 - \gamma) \|\eta_\beta \eta \varphi\|_2^2.$$

Since  $\gamma < 1$  by assumption one deduces that

$$\|H\varphi\|_2^2 \geq (1 - \gamma)^2 \|\eta_\beta \eta \varphi\|_2^2.$$

Finally since the left hand side is independent of  $\beta$  one concludes by dominated convergence that  $\eta^2\varphi \in D(H)$  and

$$\|H\varphi\|_2^2 \geq (1 - \gamma)^2 \|\eta^2\varphi\|_2^2$$

for all  $\varphi \in D(H)$ . □

Condition III of Theorem 1.1 has a different nature to the first two conditions since it is independent of  $\eta$ . It is related to the existence of an approximate identity in the  $D(\mathcal{E})$ -graph norm. In particular if there is a net  $\{\rho_\alpha\}$  with  $\rho_\alpha \in B_c(\mathcal{E})$  such that  $0 \leq \rho_\alpha \leq 1$ ,  $\rho_\alpha D(\mathcal{E}) \subseteq D(\mathcal{E})$  for all  $\alpha$  and

$$\lim_{\alpha} \|(\rho_\alpha - \mathbb{1}_X)\varphi\|_{D(\mathcal{E})} = 0 \tag{18}$$

for all  $\varphi \in D(\mathcal{E})$  then Condition III is satisfied. This follows because

$$\begin{aligned} \mathcal{E}_{\varphi^2}(\rho_\alpha) &= \mathcal{E}(\rho_\alpha \varphi) - \mathcal{E}(\varphi, \rho_\alpha^2 \varphi) \\ &= \mathcal{E}((\rho_\alpha - \mathbb{1}_X)\varphi) - \mathcal{E}(\varphi, (\rho_\alpha - \mathbb{1}_X)^2 \varphi) \end{aligned}$$

where the first step follows from Lemma 2.2 and the second by direct calculation. But  $|\mathcal{E}(\varphi, (\rho_\alpha - \mathbb{1}_X)^2 \varphi)|^2 \leq \mathcal{E}(\varphi) \mathcal{E}((\rho_\alpha - \mathbb{1}_X)^2 \varphi)$  by the Cauchy-Schwarz inequality. Then it follows from (18) and the uniform boundedness principle that there is an  $M > 0$  such that  $\mathcal{E}((\rho_\alpha - \mathbb{1}_X)^2 \varphi) \leq M \mathcal{E}((\rho_\alpha - \mathbb{1}_X)\varphi)$  for all  $\alpha$ . Therefore  $\lim_{\alpha} \mathcal{E}_{\varphi^2}(\rho_\alpha) = 0$ . In fact there is a weak converse to this statement: if Condition III of Theorem 1.1 is satisfied for all  $\varphi \in B(\mathcal{E})$  then (18) is valid for all  $\varphi \in B(\mathcal{E})$ .

### 3 Operators on domains

In this section we give the proof of Theorem 1.2. To be more precise we deduce the theorem as a corollary of Theorem 1.1. The principal difficulty is to verify Condition III of the latter theorem, the existence of a suitable approximate identity. This is the critical property used in Section 2 to extend the Rellich inequality from functions with compact support to the full domain of the submarkovian operator.

**Proof of Theorem 1.2** First the Hardy inequality (6) which is the principal assumption of the theorem can be reformulated as the  $\eta$ -Hardy inequality (2) by choosing  $\eta$  as the positive

square root of  $a_1 c_\Omega d_\Omega^{-2}$ . Then  $\eta \in W_{\text{loc}}^{1,2}(\Omega) = B_{\text{loc}}(h)$  and Condition I of Theorem 1.1 is verified. Moreover,  $h_{\varphi^2}(\eta) = (\varphi, \Gamma(\eta)\varphi)$  where the *carré du champ*  $\Gamma$  is given by  $\Gamma(\eta) = c_\Omega |\nabla \eta|^2$ . But  $4\eta^2 \Gamma(\eta) = \Gamma(\eta^2)$ . Therefore a straightforward calculation, using  $|\nabla d_\Omega| = 1$ , gives

$$\Gamma(\eta) = a_1 c_\Omega^2 d_\Omega^{-4} |1 - d_\Omega c'_\Omega / (2 c_\Omega)|^2 \leq (\nu/a_1) \eta^4$$

where  $c'_\Omega = c' \circ d_\Omega$ . The last step follows since  $\delta \wedge \delta' \leq (s c'(s)/c(s)) \leq \delta \vee \delta'$  as a consequence of the identity  $s c'(s)/c(s) = (\delta + \delta' s)/(1 + s)$ . Hence one concludes that

$$h_{\varphi^2}(\eta) = (\varphi, \Gamma(\eta)\varphi) \leq \gamma (\eta^2 \varphi, \eta^2 \varphi)$$

for all  $\varphi \in B_c(h)$  with  $\gamma = \nu/a_1$ . Then it follows from the first paragraph of the proof of Theorem 1.1 that if  $\gamma < 1$  then

$$(H\varphi, H\varphi) \geq (1 - \gamma)^2 (\eta^2 \varphi, \eta^2 \varphi) = a_1^2 (1 - \gamma)^2 (c_\Omega d_\Omega^{-2} \varphi, c_\Omega d_\Omega^{-2} \varphi)$$

for all  $\varphi \in B_c(h)$  and in particular for all  $\varphi \in C_c^\infty(\Omega)$ . Thus the Rellich inequality (7) is satisfied for all  $\varphi \in B_c(h)$  with  $a_2 = a_1^2 (1 - \gamma)^2 = (a_1 - \nu)^2$ . This is the elementary part of the proof. The difficulty lies in extending the Rellich inequality to all  $\varphi \in D(H)$ . This is achieved by the construction of an approximate identity satisfying Condition III of Theorem 1.1. The construction is an adaptation of a key idea of Agmon (see [Agm82], Chapter 1), the introduction of an alternative metric.

The Euclidean metric on  $\Omega$  is usually defined by a shortest path algorithm but it is also given by the equivalent definition

$$d(x; y) = \sup\{|\psi(x) - \psi(y)| : \psi \in W_{\text{loc}}^{1,\infty}(\Omega), |\nabla \psi| \leq 1\}.$$

The Euclidean distance from  $x \in \Omega$  to the measurable subset  $A \subseteq \Omega$  is then defined by  $d(x; A) = \inf_{y \in A} d(x; y)$  and  $d_\Omega$  is given by

$$d_\Omega(x) = \sup\{d(x; \Omega \setminus K) : K \text{ is a compact subset of } \Omega\}.$$

Now we use these definitions as the model for introducing an alternative metric.

Define the distance  $d_2(\cdot; \cdot)$  by

$$d_2(x; y) = \sup\{|\psi(x) - \psi(y)| : \psi \in W_{\text{loc}}^{1,\infty}(\Omega), d_\Omega^2 |\nabla \psi|^2 \leq 1\}.$$

Then the  $d_2$ -distance to the measurable subset  $A$  is given by  $d_2(x; A) = \inf_{y \in A} d_2(x; y)$  and the corresponding distance to the boundary by

$$d_{2;\Omega}(x) = \sup\{d_2(x; \Omega \setminus K) : K \text{ is a compact subset of } \Omega\}.$$

The functions  $d_\Omega$  and  $d_{2;\Omega}$  are Lipschitz and satisfy  $|\nabla d_\Omega|^2 = 1$  and  $d_\Omega^2 |\nabla d_{2;\Omega}|^2 = 1$  almost everywhere.

The motivation for the introduction of  $d_2(\cdot; \cdot)$  is the following.

**Lemma 3.1** *The metric space  $(\Omega, d_2(\cdot; \cdot))$  is complete.*

**Proof** It follows from Lemma A1.2 in Appendix A of Agmon's lecture notes [Agm82] that the space is complete if and only if  $d_{2,\Omega}(x) = \infty$  for one  $x \in \Omega$  or, equivalently, for all  $x \in \Omega$ . The latter equivalence is a simple application of the triangle inequality. Therefore it suffices to argue that the  $d_2$ -distance to the boundary is infinite.

Let  $\Omega_\delta = \{x \in \Omega : d_\Omega(x) < \delta\}$  for all  $\delta > 0$ . Fix  $0 < \delta_1 < \delta_2$ . Then introduce  $\psi \in W_{\text{loc}}^{1,\infty}(\Omega)$  such that

$$\psi(x) = \begin{cases} -\log(\delta_1/\delta_2) & \text{if } x \in \Omega_{\delta_1}, \\ -\log(d_\Omega(x)/\delta_2) & \text{if } x \in \overline{\Omega}_{\delta_2} \setminus \overline{\Omega}_{\delta_1}, \\ 0 & \text{if } x \in \Omega \setminus \overline{\Omega}_{\delta_2}. \end{cases}$$

Since  $d_\Omega^2 |\nabla(\log d_\Omega)|^2 = |\nabla d_\Omega|^2 \leq 1$  it follows that

$$d_2(x; y) \geq |\psi(x) - \psi(y)| = |\log(d_\Omega(x)/\delta_2) - \log(\delta_1/\delta_2)| = \log(d_\Omega(x)/\delta_1)$$

for all  $x \in \overline{\Omega}_{\delta_2} \setminus \overline{\Omega}_{\delta_1}$  and  $y \in \Omega_{\delta_1}$ . Therefore in the limit  $\delta_1 \rightarrow 0$  one deduces that  $d_{2,\Omega}(x) = \infty$  for all  $x \in \Omega$ .  $\square$

The conclusion of the foregoing argument can be rephrased as follows.

**Corollary 3.2** *For each  $m > 0$  one has  $\Omega = \{x \in \Omega : d_{2,\Omega}(x) > m\}$ .*

The completeness property of Lemma 3.1 is crucial for the verification of Condition III of Theorem 1.1. The second crucial feature is the Hardy inequality. But the following reasoning is not restricted to the forms and operators covered by Theorem 1.2. The only aspect of the Hardy inequality of relevance is the property  $c_\Omega^{1/2} d_\Omega^{-1} D(h) \subseteq L_2(\Omega)$ . We now construct a sequence of  $\rho_n$  satisfying the appropriate properties of an approximate identity by the reasoning of Agmon [Agm82] in the proof of his Theorem 1.5.

**Proposition 3.3** *Assume that  $h$  satisfies the Hardy inequality (6). Then there exists a sequence  $\rho_n \in D(h)$  with compact support such that  $0 \leq \rho_n \leq 1$  and*

$$\lim_{n \rightarrow \infty} \|(\rho_n - \mathbf{1}_\Omega)\varphi\|_2 = 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} h_{\varphi^2}(\rho_n) = 0$$

for all  $\varphi \in D(h)$ .

**Proof** First let  $\{\Omega_n\}_{n \geq 1}$  be a compact exhaustion sequence of  $\Omega$ , i.e. the  $\Omega_n$  are relatively compact open subsets of  $\Omega$  with  $\overline{\Omega}_n \subset \Omega_{n+1}$  such that  $\Omega = \bigcup_{n \geq 1} \Omega_n$ . Secondly, fix  $m > 0$  and define  $\rho$  by  $\rho(t) = (t/m) \wedge 1$ . Further define  $\rho_n$  by  $\rho_n(x) = \rho(d_2(x; \Omega \setminus \Omega_n))$ . Then

$$\rho_n(x) = \begin{cases} 0 & \text{if } x \in \Omega \setminus \Omega_n, \\ d_2(x; \Omega \setminus \Omega_n)/m & \text{if } x \in \Omega_n \text{ and } d_2(x; \Omega \setminus \Omega_n) \leq m, \\ 1 & \text{if } x \in \Omega_n \text{ and } d_2(x; \Omega \setminus \Omega_n) > m. \end{cases}$$

It follows immediately from this definition that the  $\rho_n$  have compact support. In particular  $\text{supp } \rho_n \subseteq \overline{\Omega}_n$ . Moreover,  $0 \leq \rho_n \leq 1$ . In addition the pointwise limit of the  $\rho_n$  as  $n \rightarrow \infty$  is equal to the identity on the set of  $x$  for which  $d_{2,\Omega}(x) > m$ . But this set is equal to  $\Omega$  by Corollary 3.2. Thus the  $\rho_n$  converge pointwise to  $\mathbf{1}_\Omega$  as  $n \rightarrow \infty$ . Hence they also converge

to the identity strongly on  $L_2(\Omega)$ . Finally the  $\rho_n$  are Lipschitz continuous. This follows by first noting that

$$|\rho(s) - \rho(t)| \leq m^{-1}|s - t|.$$

Therefore

$$\begin{aligned} |\rho_n(x) - \rho_n(y)| &\leq m^{-1}|d_2(x; \Omega \setminus \Omega_n) - d_2(y; \Omega \setminus \Omega_n)| \\ &\leq m^{-1}d_2(x; y) \leq C m^{-1}|x - y|. \end{aligned}$$

The second inequality follows from the triangle inequality and is valid for all  $x, y \in \Omega$ . The third inequality follows from the definition of  $d_2(\cdot; \cdot)$  and is valid locally, i.e. it is valid in a neighbourhood of each fixed point  $x_0 \in \Omega$  with the value of  $C$  depending on  $x_0$ . Consequently, since the  $\rho_n$  are Lipschitz functions and  $|\rho_n(x) - \rho_n(y)| \leq m^{-1}d_2(x; y)$  for all  $x, y \in \Omega$  it follows from the eikonal inequality (see [Agm82], Theorem 1.4 (ii)) that

$$d_\Omega^2 |\nabla \rho_n|^2 \leq m^{-2}.$$

This is the critical inequality since it gives

$$\Gamma(\rho_n) = c_\Omega |\nabla \rho_n|^2 \leq m^{-2} c_\Omega d_\Omega^{-2}.$$

This estimate corresponds to (1.23) in Agmon's notes. Next note that the derivatives of  $\rho_n$  have support in the set  $\Omega_{m,n} = \{x \in \Omega_n : d_2(x; \Omega \setminus \Omega_n) \leq m\}$  and  $|\Omega_{m,n}| \rightarrow 0$  as  $n \rightarrow \infty$  by Corollary 3.2. Moreover, if  $K$  is a compact subset of  $\Omega$  then  $K \subset \Omega_n$  for all sufficiently large  $n$ . Up to this point we have not used the Hardy inequality (6). But it follows from this inequality that if  $\varphi \in D(h)$  then  $\varphi \in D(c_\Omega^{1/2} d_\Omega^{-1})$ . Therefore  $\psi = c_\Omega^{1/2} d_\Omega^{-1} \varphi \in L_2(\Omega)$ . Hence

$$\int_\Omega \Gamma(\rho_n) |\varphi|^2 \leq m^{-2} \int_{\Omega_{m,n}} c_\Omega d_\Omega^{-2} |\varphi|^2 = m^{-2} \int_{\Omega_{m,n}} |\psi|^2.$$

Since  $\psi \in L_2(\Omega)$  it follows directly from this estimate that

$$\lim_{n \rightarrow \infty} \int_\Omega \Gamma(\rho_n) |\varphi|^2 = 0.$$

Thus, in the earlier notation,  $h_{\varphi^2}(\rho_n) \rightarrow 0$  as  $n \rightarrow \infty$ . □

**Proof of Theorem 1.2 continued** The proof of the theorem is now a corollary of Theorem 1.1. Condition I of the theorem is valid by assumption of the Hardy inequality (6), Condition II was verified at the beginning of the proof with  $\gamma = \nu/a_1$  and Condition III now follows from Proposition 3.3. Therefore the Rellich inequality (7) follows for all  $\varphi \in D(H)$  with  $a_2 = (a_1 - \nu)^2$  whenever  $\nu < a_1$ . □

## 4 Applications and illustrations

In this section we give two illustrations of Theorems 1.1 and 1.2. First we give a direct application of the latter theorem with  $\Omega = \mathbf{R}^d \setminus \{0\}$ . The application requires establishing the validity of the Hardy inequality (6), calculating the Hardy constant  $a_1$  and verifying the condition  $a < \nu_1$ .

**Example 4.1** Let  $\Omega = \mathbf{R}^d \setminus \{0\}$ . Then  $\partial\Omega = \{0\}$  and  $d_\Omega(x) = |x|$ . The Dirichlet form  $h$ , submarkovian operator  $H$ , coefficient function  $c$  etc. are defined in Theorem 1.2. In particular  $c(s) = s^\delta(1+s)^{\delta'-\delta}$  with  $\delta, \delta' \geq 0$  and  $c_\Omega(x) = c(|x|)$ .

**Observation 4.2** If  $d + (\delta \wedge \delta') - 2 > 0$  then the Hardy inequality

$$h(\varphi) \geq a_1 (c_\Omega^{1/2} d_\Omega^{-1} \varphi, c_\Omega^{1/2} d_\Omega^{-1} \varphi) \quad (19)$$

is valid for all  $\varphi \in D(h)$  with  $a_1 = (d + (\delta \wedge \delta') - 2)^2/4$ . The value of  $a_1$  is optimal.

**Proof** First note that

$$\begin{aligned} h(\varphi) - \lambda (\varphi, (\operatorname{div} c_\Omega \chi) \varphi) + \lambda^2 (\varphi, c_\Omega \chi^2 \varphi) \\ = ((\nabla + \lambda \chi) \varphi, c_\Omega (\nabla + \lambda \chi) \varphi) \geq 0 \end{aligned} \quad (20)$$

for all  $\varphi \in C_c^\infty(\Omega)$  and  $\chi = (\chi_1, \dots, \chi_d)$  with  $\chi_k \in W_{\text{loc}}^{1,\infty}(\Omega)$ . Now choose  $\chi = (\nabla d_\Omega) d_\Omega^{-1}$ . Thus  $\chi(x) = x|x|^{-2}$ . Since  $|\nabla d_\Omega| = 1$  it follows that  $c_\Omega \chi^2 = c_\Omega d_\Omega^{-2}$ . Moreover,

$$\operatorname{div}(c_\Omega \chi) = (d - 2 + c'_\Omega d_\Omega / c_\Omega) c_\Omega d_\Omega^{-2} \geq (d + (\delta \wedge \delta') - 2) c_\Omega d_\Omega^{-2}$$

since  $c'(s) s / c(s) \geq \delta \wedge \delta'$ . Then one deduces from (20) that

$$h(\varphi) - 2\lambda b (\varphi, c_\Omega d_\Omega^{-2} \varphi) + \lambda^2 (\varphi, c_\Omega d_\Omega^{-2} \varphi) \geq 0 \quad (21)$$

for all  $\lambda > 0$  where  $b = (d + (\delta \wedge \delta') - 2)/2$ . Then if  $b > 0$  one can choose  $\lambda = b$  and the Hardy inequality follows, with  $a_1 = b^2 = (d + (\delta \wedge \delta') - 2)^2/4$ , for all  $\varphi \in C_c^\infty(\Omega)$  and then by continuity for all  $\varphi \in D(h)$ .

The optimality of  $a_1$  follows by variation a standard argument (see, for example, [BEL15] Chapter 1). First let  $a$  denote the optimal value of the constant for a Hardy inequality of the form (19). Secondly, set  $a(\delta) = ((d + \delta - 2)/2)^2$ . Then it follows from (19) that  $a \geq a(\delta \wedge \delta')$ . Therefore it suffices to prove the identical upper bound. But then it is sufficient to prove that  $a(\delta)$  and  $a(\delta')$  are both upper bounds since  $a(\delta \wedge \delta') = a(\delta) \wedge a(\delta')$ . The first bound is established by an estimate at the origin and the second by a similar estimate at infinity.

The estimate at the origin is obtained by examining functions  $\varphi_\alpha = d_\Omega^{-\alpha} \xi$ ,  $\alpha > 0$ , where  $\xi$  has support in a small neighbourhood of the origin. Then  $c_\Omega d_\Omega^{-2} |\varphi_\alpha|^2$  is integrable if  $\alpha < (d + \delta - 2)/2$ . Choosing  $\alpha = (d + \delta - 2 - \varepsilon)/2$  with  $\varepsilon > 0$  small one can arrange that  $h(\varphi_\alpha) / \|c_\Omega^{1/2} d_\Omega^{-1} \varphi_\alpha\|_2^2 \leq \alpha^2$  and in the limit  $\varepsilon \rightarrow 0$  one concludes that  $a \leq a(\delta)$ . Here the property  $\lim_{s \rightarrow 0} c(s) s^{-\delta} = 1$  is important. The estimate at infinity is similar. One now chooses  $\varphi_\alpha$  with support in the complement of a large ball centred at the origin and equal to  $d_\Omega^{-\alpha}$  outside a larger ball. Then  $c_\Omega d_\Omega^{-2} |\varphi_\alpha|^2$  is integrable if  $\alpha > (d + \delta' - 2)/2$ . So choosing  $\alpha = (d + \delta' - 2 + \varepsilon)/2$  and proceeding as in the local approximation one deduces that  $a \leq a(\delta')$ . Here the property  $\lim_{s \rightarrow \infty} c(s) s^{-\delta'} = 1$  is crucial.  $\square$

Now one can apply Theorem 1.2 to obtain the Rellich inequality. It suffices to compute  $\nu$  and verify that  $\nu < a_1$ . There are two distinct cases.

**Observation 4.3** Assume  $\delta + \delta' \leq 4$ . If  $d + 2(\delta \wedge \delta') - 4 > 0$  then the Rellich inequality

$$(H\varphi, H\varphi) \geq a_2 (c_\Omega d_\Omega^{-2} \varphi, c_\Omega d_\Omega^{-2} \varphi) \quad (22)$$

is valid for all  $\varphi \in D(H)$  with  $a_2 = d^2(d + 2(\delta \wedge \delta') - 4)^2/16$ . The value of  $a_2$  is optimal.

**Proof** First consider the case  $\delta, \delta' \in [0, 2]$ . Then

$$\nu = \sup\{|1 - t/2|^2 : \delta \wedge \delta' \leq t \leq \delta \vee \delta'\} = (1 - (\delta \wedge \delta')/2)^2$$

and  $\nu < a_1$  if and only if  $1 - (\delta \wedge \delta')/2 < (d + (\delta \wedge \delta') - 2)/2$  or, equivalently,  $d + 2(\delta \wedge \delta') - 4 > 0$ . But this implies the condition  $d + (\delta \wedge \delta') - 2 > 0$  necessary for the Hardy inequality (19). Therefore one deduces from Theorem 1.2 that the Rellich inequality (7) is valid with constant  $a_2 = (a_1 - \nu)^2$  which is easily calculated to be the value given in the observation.

Secondly, assume  $\delta \in [0, 2]$  but  $\delta' \geq 2$ . Then  $0 \leq \delta'/2 - 1 \leq 1 - \delta/2$  since  $\delta + \delta' \leq 4$ . Therefore

$$\nu = \sup\{|1 - t/2|^2 : \delta \leq t \leq \delta'\} = |1 - \delta/2|^2 = |1 - (\delta \wedge \delta')/2|^2.$$

Hence the Rellich inequality (22) is again valid with the same value of  $a_2$ .

Thirdly, if  $\delta' \in [0, 2]$  but  $\delta \geq 2$  one reaches the same conclusion by interchanging  $\delta$  and  $\delta'$  in the last argument.

Therefore the observation is established for all  $\delta, \delta' \geq 0$  with  $\delta + \delta' \leq 4$ .

The optimality of  $a_2$  follows by a reasoning similar to the Hardy case. One again needs separate arguments at the origin and at infinity.  $\square$

**Observation 4.4** Assume  $\delta + \delta' \geq 4$ . If  $d - |\delta - \delta'| > 0$  then the Rellich inequality (22) is valid with  $a_2 = (d - |\delta - \delta'|)^2(d + \delta + \delta' - 4)^2/16$ .

**Proof** First assume  $\delta, \delta' \geq 2$ . Then  $\nu = (1 - (\delta \vee \delta')/2)^2$ . Therefore  $a_1 > \nu$  if and only if  $d > (\delta \vee \delta') - (\delta \wedge \delta') = |\delta - \delta'|$ . This condition also implies  $d + (\delta \wedge \delta') \geq (\delta \vee \delta') \geq 2$ . Therefore the Hardy inequality (19) is valid. Moreover,  $a_1 - \nu = (d + \delta + \delta' - 4)(d - |\delta - \delta'|)/4$  and one deduces from Theorem 3 that the Rellich inequality (22) is valid with  $a_2 = (a_1 - \nu)^2$  whenever  $d > |\delta - \delta'|$ .

Secondly, assume  $\delta \leq 2$  and  $\delta' \geq 2$ . Since  $\delta + \delta' \geq 4$  it follows that  $1 - \delta/2 \leq \delta'/2 - 1$ . Therefore  $\nu = (1 - \delta'/2)^2 = (1 - (\delta \vee \delta')/2)^2$ . Then the observation follows again.

The final case  $\delta \geq 2$  and  $\delta' \leq 2$  now follows from the second case by interchanging  $\delta$  and  $\delta'$ .  $\square$

There is one question left over in this discussion of Example 4.1, the optimality of the value of  $a_2$  in Observation 4.4. It does follow from the argument in the proof of Observation 4.3 that the value in the case  $\delta + \delta' \geq 4$  is less than or equal to the value in the case  $\delta + \delta' \leq 4$ . Moreover the two values are equal if and only if  $\delta = \delta'$  or  $\delta + \delta' = 4$ . This follows by noting that

$$d(d + 2(\delta \wedge \delta') - 4) - (d - |\delta - \delta'|)(d + \delta + \delta' - 4) = |\delta - \delta'|(\delta + \delta' - 4).$$

Therefore the optimal value in the case  $\delta + \delta' \geq 4$  is generally strictly smaller than the value  $d^2(d + \delta \wedge \delta' - 4)^2/16$ .

As a second illustration of the foregoing techniques we consider a general class of operators of Grushin type which are related to the classic situation described in Example 4.1. These operators differ somewhat from the standard Grushin operators. Many of their properties, e.g. Gaussian kernel bounds, Poincaré inequalities, etc., were previously established in [RS08a] [RS08b] [RS14]. Although the Grushin operators are not directly covered by Theorem 1.2 similar conclusions can be drawn by a slight modification of the proof of Theorem 1.1.



**Example 4.5** Let  $\Omega = (\mathbf{R}^{d_1} \setminus \{0\}) \times \mathbf{R}^{d_2}$  and set  $x = (x_1, x_2)$  with  $x_1 \in \mathbf{R}^{d_1}$  and  $x_2 \in \mathbf{R}^{d_2}$ . Then  $\partial\Omega = \{x = (0, x_2) : x_2 \in \mathbf{R}^{d_2}\}$  and  $d_\Omega(x) = |x_1|$ . Next define the Dirichlet form  $h$  on  $L_2(\Omega) = L_2(\mathbf{R}^{d_1} \setminus \{0\}) \otimes L_2(\mathbf{R}^{d_2})$  as the closure of the form

$$\varphi \in C_c^\infty(\Omega) \mapsto h(\varphi) = (\nabla_{x_1} \varphi, c_\Omega \nabla_{x_1} \varphi) + (\nabla_{x_2} \varphi, b \nabla_{x_2} \varphi) \quad (23)$$

where  $c_\Omega = c \circ d_\Omega$  with  $c$  again the function defined in Theorem 1.2 and  $b$  the operator of multiplication by a positive bounded function of the first variable  $x_1$ . Thus the coefficient  $c_\Omega$  of the first form on the right,  $h^{(1)}$ , and the coefficient  $b$  of the second form,  $h^{(2)}$ , are both independent of  $x_2$ .

The forms  $h^{(1)}$  and  $h^{(2)}$  are both closable on  $L_2(\Omega)$  and their closures are Dirichlet forms. The submarkovian operator  $H_1$  associated with  $h^{(1)}$  is the tensor product of an operator  $\tilde{H}_1$  which acts on the first component  $L_2(\mathbf{R}^{d_1} \setminus \{0\})$  of the tensor product space and the identity operator  $\mathbb{1}_2$  on the second component  $L_2(\mathbf{R}^{d_2})$ . But  $\tilde{H}_1$  can be identified as the operator analyzed in Example 4.1. Therefore it satisfies the Hardy inequality (19) on  $L_2(\mathbf{R}^{d_1} \setminus \{0\})$ , with  $d$  replaced by  $d_1$ . Now since  $h \geq h^{(1)}$  the form  $h$  satisfies the corresponding Hardy inequality on  $L_2(\Omega)$ . Explicitly one has the following.

**Observation 4.6** *If  $d_1 + (\delta \wedge \delta') - 2 > 0$  then*

$$h(\varphi) \geq a_1 (c_\Omega^{1/2} d_\Omega^{-1} \varphi, c_\Omega^{1/2} d_\Omega^{-1} \varphi) \quad (24)$$

for all  $\varphi \in D(h)$  with  $a_1 = (d_1 + (\delta \wedge \delta') - 2)^2/4$ . The value of  $a_1$  is optimal.

**Proof** It only remains to prove that the constant  $a_1$  is optimal. But if  $\tilde{a}$  is the optimal constant then clearly  $\tilde{a} \geq a_1$  and it suffices to prove the converse bound.

The optimal value  $\tilde{a}$  is given by

$$\begin{aligned} \tilde{a} &= \inf \{h(\varphi) / \|c_\Omega^{1/2} d_\Omega^{-1} \varphi\|_2^2 : \varphi \in D(h)\} \\ &\leq \inf \{h(\psi \chi) / (\|c_\Omega^{1/2} d_\Omega^{-1} \psi\|_2^2 \|\chi\|_2^2) : \psi \in D(h^{(1)}), \chi \in C_c^\infty(\mathbf{R}^{d_2})\} \end{aligned}$$

where we have slightly abused notation by not distinguishing between the  $L_2$ -norms on the two components in the tensor product space. It follows, however, from the product structure that

$$h(\psi \chi) = h^{(1)}(\psi) \|\chi\|_2^2 + (b \psi, \psi) \|\nabla_{x_2} \chi\|_2^2.$$

Next replace  $\chi$  by  $\chi_\lambda : \chi_\lambda(x_2) = \lambda^{d_2/2} \chi(\lambda x_2)$ . Since  $\|\chi_\lambda\|_2 = \|\chi\|_2$  and  $\|\nabla_{x_2} \chi_\lambda\|_2 = \lambda \|\nabla_{x_2} \chi\|_2$  it follows immediately that

$$\lim_{\lambda \rightarrow 0} h(\psi \chi_\lambda) / (\|c_\Omega^{1/2} d_\Omega^{-1} \psi\|_2^2 \|\chi_\lambda\|_2^2) = h^{(1)}(\psi) / \|c_\Omega^{1/2} d_\Omega^{-1} \psi\|_2^2 = a_1$$

where the last identification follows from Example 4.1. Therefore  $\tilde{a} = a_1$ .  $\square$ .

Next we argue that the submarkovian operator  $H$  corresponding to the Grushin form  $h$  satisfies a Rellich inequality. Theorem 1.2 is not directly applicable as the Grushin form has the second component  $h^{(2)}$ . But in fact the Rellich inequality is independent of  $h^{(2)}$ . This is somewhat surprising but can be understood by revisiting the proof of Theorem 1.1. First we state the result. There are again two distinct regimes.

**Observation 4.7** Let  $H$  be the submarkovian operator on  $L_2(\Omega)$  corresponding to the Grushin form (23).

Assume  $\delta + \delta' \leq 4$ . If  $d_1 + 2(\delta \wedge \delta') - 4 > 0$  then the Rellich inequality

$$(H\varphi, H\varphi) \geq a_2 (c_\Omega d_\Omega^{-2} \varphi, c_\Omega d_\Omega^{-2} \varphi) \quad (25)$$

is valid for all  $\varphi \in D(H)$  with  $a_2 = d_1^2(d_1 + 2(\delta \wedge \delta') - 4)^2/16$ . The value of  $a_2$  is optimal.

Alternatively assume  $\delta + \delta' \geq 4$ . If  $d - |\delta - \delta'| > 0$  then the Rellich inequality (25) is valid with  $a_2 = (d_1 - |\delta - \delta'|)^2(d_1 + \delta + \delta' - 4)^2/16$ .

**Proof** The Rellich inequality follows from the Hardy inequality of Observation 4.6 by a modification of the proof of Theorem 1.1 with  $\mathcal{E} = h = h^{(1)} + h^{(2)}$ . The idea is to show that the main estimates of the proof are all independent of  $h^{(2)}$ .

First  $h$  satisfies the  $\eta$ -Hardy inequality with  $\eta = c_\Omega^{1/2} d_\Omega^{-1}$  by Observation 4.6 and  $\eta$  is a function of  $x_1$ . Secondly, the form  $h$  has a *carré du champ*  $\Gamma$  given by

$$\Gamma(\varphi) = c_\Omega |\nabla_{x_1} \varphi|^2 + b |\nabla_{x_2} \varphi|^2.$$

Then since  $\eta$  is independent of  $x_2$  one has  $\Gamma(\eta) = c_\Omega |\nabla_{x_1} \eta|^2$ . Thirdly it follows from the proof of Theorem 1.2 that there exists an approximate identity  $\rho_\alpha$  satisfying Condition III of Theorem 1.1 on  $L_2(\mathbf{R}^{d_1} \setminus \{0\})$ . Therefore we can construct the bounded approximants  $\eta_{\alpha,\beta}$  as in the proof of the latter theorem and these remain functions of  $x_1$ .

The key identity (13) in the proof of Theorem 1.1 now takes the form

$$(H\varphi, \eta_{\alpha,\beta}^2 \varphi_\varepsilon) = h(\eta_{\alpha,\beta} \varphi_\varepsilon) - h_{\varphi_\varepsilon^2}(\eta_{\alpha,\beta}) + h(\varphi - \varphi_\varepsilon, \eta_{\alpha,\beta}^2 \varphi_\varepsilon)$$

with  $\varphi_\varepsilon$  again the bounded approximate to  $\varphi \in D(H)$ . But  $h \geq h^{(1)}$  and

$$h_{\varphi_\varepsilon^2}(\eta_{\alpha,\beta}) = (\varphi, \Gamma(\eta_{\alpha,\beta}) \varphi) = h_{\varphi_\varepsilon^2}^{(1)}(\eta_{\alpha,\beta})$$

because  $\Gamma(\eta_{\alpha,\beta})$  is independent of  $x_2$ . Therefore one obtains the estimate

$$\begin{aligned} \|H\varphi\|_2 \|\eta_{\alpha,\beta}^2 \varphi_\varepsilon\|_2 &\geq h^{(1)}(\eta_{\alpha,\beta} \varphi_\varepsilon) - h_{\varphi_\varepsilon^2}^{(1)}(\eta_{\alpha,\beta}) + h(\varphi - \varphi_\varepsilon, \eta_{\alpha,\beta}^2 \varphi_\varepsilon) \\ &\geq h^{(1)}(\eta_{\alpha,\beta} \varphi_\varepsilon) - h_{\varphi_\varepsilon^2}^{(1)}(\eta_{\alpha,\beta}) - h^{(1)}(\varphi - \varphi_\varepsilon)^{1/2} h^{(1)}(\eta_{\alpha,\beta}^2 \varphi_\varepsilon)^{1/2} \\ &\quad - h^{(2)}(\varphi - \varphi_\varepsilon)^{1/2} h^{(2)}(\eta_{\alpha,\beta}^2 \varphi_\varepsilon)^{1/2} \end{aligned}$$

and the only dependence on  $h^{(2)}$  is in the last term which converges to zero as  $\varepsilon \rightarrow 0$ . This last point depends on the uniform bound  $h^{(2)}(\eta_{\alpha,\beta}^2 \varphi_\varepsilon) \leq \|\eta_\beta\|_\infty^4 h^{(2)}(\varphi_\varepsilon) \leq \|\eta_\beta\|_\infty^4 h^{(2)}(\varphi)$ . Therefore one can now repeat the proof of Theorem 1.1 following the identity (13) to obtain a Rellich inequality which is totally independent of  $h^{(2)}$ . The Rellich inequality is determined by  $h^{(1)}$ . The end result is (25) with  $a_2 = (a_1 - \nu)^2$  where  $a_1$  is the Hardy constant given in Observation 4.6 and  $\nu$  is again given by  $\sup\{|1 - t/2|^2 : \delta \wedge \delta' \leq t \leq \delta \vee \delta'\}$ . The calculation is a repetition of that given in Example 4.1. The only difference is that  $h^{(1)}$  is now a form on the first component of the tensor product space  $L_2(\mathbf{R}^{d_1} \setminus \{0\}) \otimes L_2(\mathbf{R}^{d_2})$  but this makes no essential difference.

The only remaining point to establish is the optimality of  $a_2$  in the case that  $\delta + \delta' \leq 4$ . But this follows by a slight generalization of the argument used to prove the optimality of  $a_1$  in Observation 4.6. Now, however, one uses the tensor product structure to note that

$$\|H(\psi \chi)\|_2^2 = \|H_1 \psi\|_2^2 \|\chi\|_2^2 + 2(b\psi, H_1 \psi) \|\nabla_{x_2} \chi\|_2^2 + \|b\psi\|_2^2 \|\Delta_{x_2} \chi\|_2^2$$

with  $\psi \in L_2(\mathbf{R}^{d_1} \setminus \{0\})$  and  $\chi \in L_2(\mathbf{R}^{d_2})$ . □

# Acknowledgements

The author is indebted to Juha Lehrbäck for a continuing, informative, correspondence on the Hardy and Rellich inequalities together with several specific suggestions about the current material. His advice has been very helpful.

# References

- [Agm82] AGMON, S., *Lectures on exponential decay of solutions of second-order elliptic equations*. Mathematical Notes 29. Princeton University Press, Princeton, 1982.
- [AH05] ARIYOSHI, T., and HINO, M., Small-time asymptotic estimates in local Dirichlet spaces. *Elec. J. Prob.* **10** (2005), 1236–1259.
- [And75] ANDERSSON, L.-E., On the representation of Dirichlet forms. *Ann. Inst. Fourier* **25** (1975), 11–25.
- [BEL15] BALINSKY, A. A., EVANS, W. D., and LEWIS, R. L., *The Analysis and Geometry of Hardy’s Inequality*. Universitext. Springer, New York, 2015.
- [BH91] BOULEAU, N., and HIRSCH, F., *Dirichlet forms and analysis on Wiener space*, vol. 14 of de Gruyter Studies in Mathematics. Walter de Gruyter & Co., Berlin, 1991.
- [DS83] DRIESSLER, W., and SUMMERS, S. J., On commutators and selfadjointness. *Lett. Math. Phys.* **7** (1983), 319–326.
- [ER09] ELST, A. F. M. TER, and ROBINSON, D. W., Uniform subellipticity. *J. Operator Theory* **62** (2009), 125–149.
- [ERSZ06] ELST, A. F. M. TER, ROBINSON, D. W., SIKORA, A., and ZHU, Y., Dirichlet forms and degenerate elliptic operators. In KOELINK, E., NEERVEN, J. VAN, PAGTER, B. DE, and SWEERS, G., eds., *Partial Differential Equations and Functional Analysis*, vol. 168 of Operator Theory: Advances and Applications. Birkhäuser, 2006, 73–95. Philippe Clement Festschrift.
- [Far75] FARIS, W. G., *Self-adjoint operators*. Lect. Notes in Math. 433. Springer-Verlag, Berlin etc., 1975.
- [FOT94] FUKUSHIMA, M., OSHIMA, Y., and TAKEDA, M., *Dirichlet forms and symmetric Markov processes*, vol. 19 of de Gruyter Studies in Mathematics. Walter de Gruyter & Co., Berlin, 1994.
- [FP83] FEFFERMAN, C., and PHONG, D. H., Subelliptic eigenvalue problems. In *Conference on harmonic analysis in honor of Antoni Zygmund*, Wadsworth Math. Ser., 590–606. Wadsworth, Belmont, CA, 1983.
- [GJ72] GLIMM, J., and JAFFE, A., The  $\lambda\phi_2^4$  quantum field theory without cutoffs. IV. Perturbations of the Hamiltonian. *J. Math. Phys.* **13** (1972), 1568–1584.

- [GJ81] ———, *Quantum Physics. A Functional integral point of view*. Springer-Verlag, New York etc., 1981.
- [Gri03] GRILLO, G., Hardy and Rellich-Type Inequalities for Metrics Defined by Vector Fields. *Pot. Anal.* **18** (2003), 187–217.
- [MR92] MA, Z. M., and RÖCKNER, M., *Introduction to the theory of (non symmetric) Dirichlet Forms*. Universitext. Springer-Verlag, Berlin etc., 1992.
- [Rot76] ROTH, J.-P., Formule de représentation et troncature des formes de Dirichlet sur  $\mathbf{R}^m$ . In *Séminaire de Théorie du Potentiel de Paris, No. 2*, Lect. Notes in Math. 563, 260–274. Springer Verlag, Berlin, 1976.
- [RS75] REED, M., and SIMON, B., *Methods of modern mathematical physics II. Fourier analysis, self-adjointness*. Academic Press, New York etc., 1975.
- [RS08a] ROBINSON, D. W., and SIKORA, A., Analysis of degenerate elliptic operators of Grušin type. *Math. Z.* **260** (2008), 475–508.
- [RS08b] ———, Analysis of degenerate elliptic operators of Grušin type, 2008. arXiv:1305.7002 math.AP.
- [RS11] ———, Markov uniqueness of degenerate elliptic operators. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* **10** (2011), 731–759.
- [RS14] ———, The limitations of the Poincaré inequality for Grušin type operators. *J. Evol. Equ.* **14** (2014), 535–563.
- [War14] WARD, A. D., *On Essential Self-adjointness, Confining Potentials and the  $L_p$ -Hardy inequality*. PhD thesis, Massey University, Albany, New Zealand, 2014. <http://hdl.handle.net/10179/5941>.